Slide 1 of 18

Title Slide: Getting Started with the International Activities Program Data

Slide 2 of 18

This module provides users with an overview of the data collected through the IAP studies that are available for researchers to analyze. The studies which will be covered in this module are Progress in International Reading Literacy Study, or PIRLS, Trends in International Mathematics and Science Study, or TIMSS, and Program for International Student Assessment, or PISA.

This module also describes which resources are available to learn more about the studies, the data, and the data files. Information presented in this module is key to understanding some of the more detailed information presented in subsequent IAP modules. For this reason, users who are planning to proceed through the subsequent IAP modules and use IAP study data for analytic purposes are strongly encouraged to complete this module first.

Slide 3 of 18

In all IAP studies there are separate data files by data type—for example, student data files, school data files, etc. PIRLS and TIMSS data are stored separately by data type and education system. In both PIRLS and TIMSS, education systems include countries as a whole, as well as subnational entities such as U.S. states, Canadian provinces, and special administrative regions.

For PISA, the data are stored separately by data type, but all education systems are in the same file. Participating entities within PISA are also referred to as education systems and include Organization for Economic Cooperation and Development countries (or OECD countries), some non-OECD countries, as well as subnational entities such as U.S. states and special administrative regions. It is important to note that participating education systems vary across IAP studies and within IAP studies from one administration to the next.

Slide 4 of 18

Across the three main IAP studies, there are two basic types of data files. The first type is the background files, which include student contextual data files, school contextual data files, etc. PIRLS and PISA include home data files derived from a parent questionnaire, though U.S. data are not included because the United States has not administered these parent questionnaires. The second type is the achievement or cognitive files. While test scores are included in the student background data files, more detailed assessment information—that is, student responses to assessment items—can be found in the achievement or cognitive files. For PIRLS and TIMSS, the separate files by data type and education system can be merged using the IEA International Database Analyzer, or IDB Analyzer. For PISA, if you need to merge files (for example, merging

the school data file with the student data file), you will need to do this yourself (such as using the Data – Merge Files function in SPSS) before using the IDB Analyzer.

Slide 5 of 18

The student background files contain the responses to the student questionnaire that asks students about their: attitudes towards the subject assessed; demographic backgrounds; school experiences; career and educational goals; and their beliefs, values, and attitudes. The student background files also contain achievement data, which include sampled students' scale and subscale scores.

Slide 6 of 18

The teacher background files contain responses to a teacher questionnaire that asks teachers about their: background and education; instructional practices; pedagogical beliefs; and their beliefs about the subject tested. Teacher background files are available for PIRLS and TIMSS, as well as for PISA starting with the 2015 administration.

Slide 7 of 18

Analysts of PIRLS and TIMSS data should use caution when using the teacher background files. The teacher data provided describes teachers who teach a representative sample of students. Therefore, statements can be made about the percent of students taught by teachers, and NOT the percent of teachers who do one thing or another.

To conduct analyses on what teachers do or do not do, you would need a sample of teachers rather than teachers of sampled students. Such data are collected by the Teaching and Learning International Survey (TALIS). TIMSS and PIRLS only collect data from teachers of sampled students to provide contextual data for students. Schools and students within those schools are selected for participation in the study, and then the teachers who teach those students are selected. This does not provide a representative sample of teachers within an education system, as larger classes have a greater probability of being selected than smaller classes. Thus, the teachers of the smaller classes will not be selected at the same rate as teachers from larger classes.

PISA does not collect any teacher data because it does not specifically assess what is taught in the classroom, but rather how well students can apply what they know to real-world problems.

Slide 8 of 18

Analysts of PIRLS, TIMSS, and PISA data should use caution when using the teacher and school background files.

To conduct analyses on what teachers do or do not do, you would need a sample of teachers rather than teachers of sampled students. Such data are collected by the

Teaching and Learning International Survey (TALIS). PIRLS, TIMSS, and PISA only collect data from teachers of sampled students to provide contextual data for students. Schools and students within those schools are selected for participation in the study, and then the teachers who teach those students are selected. This does not provide a representative sample of teachers within an education system, as larger classes have a greater probability of being selected than smaller classes. Thus, the teachers of the smaller classes will not be selected at the same rate as teachers from larger classes.

Therefore, analyses with teacher data should be made with students as the units of analysis and reported in terms of students who are taught by teachers with a particular attribute. Teacher data must be analyzed by linking the students to their teachers. In PIRLS and TIMSS, this is done by using the student–teacher linkage data files. Statements, then, can be made about the percent of students taught by teachers, and NOT the percent of teachers who do one thing or another.

Because PIRLS, TIMSS, and PISA have representative samples of schools, it is possible to compute reasonable statistics with schools as units of analysis. However, the school samples were designed to optimize the student samples and the student-level estimates. For this reason, it is preferable to analyze school-level variables as attributes of the students, rather than as elements in their own right. Therefore, analyzing school data should be done by linking the students to their schools, that is, merging the student and school background data files.

Slide 9 of 18

The achievement test files, which are called the cognitive items files in PISA, contain students' responses to the assessment items administered. PIRLS and TIMSS include unscored achievement files, while PISA includes both scored and unscored cognitive data files. All students are administered both multiple-choice and constructed-response items. It is important to remember that each student only receives a fraction of the items in the test pool.

Slide 10 of 18

The TIMSS and PIRLS data almanacs and the PISA compendia are files that display previously computed, weighted summary statistics, by grade level, for each participating education system, on each variable included in the background questionnaires (except the curriculum questionnaire), as well as in the test booklets. There is one almanac or compendia for each background instrument and subject area. The almanac files are in Microsoft Word and Adobe Acrobat files, while the compendia files are in Excel.

The almanacs and compendia are valuable for several reasons. First, they are useful for checking or refining your hypothesis by reviewing relevant results that have already been computed. Second, they may also be used to check the results of your own computations. Finally, you can search them for variables that you might be interested in investigating by using the Find tool in Microsoft Word and Excel. For example, if you are interested in how access to books might be related to student's knowledge of a subject,

you can enter “books” into the search engine and find any background questions that asked about books.

Links to each of these resources, by IAP study, are provided in the module titled, Data Collected Through the IAP Studies,’ which can be accessed by clicking on the corresponding underlined screen text.

Slide 11 of 18

For each categorical variable in the background questionnaires and test booklets, the data almanacs and compendia provide data on the sample size, weighted percent of responses by category, percent of missing responses, and mean achievement by response category.

For each continuous variable in the background questionnaires and test booklets, the data almanacs provide data on the sample size, percent of missing responses, and points in the distribution, which include the mean, mode, minimum, maximum, and selected percentiles.

Slide 12 of 18

This is an example item from the TIMSS 2011 8th-grade almanac for mathematics. It shows results, by education system, for the item asking, “Are you a girl or a boy?”

Slide 13 of 18

Also accompanying the PIRLS, TIMSS, and PISA data files are codebooks. These are guides that describe the layout of the data in the data file and document the values associated with the answer options provided for each survey question. This is an example item from the PISA 2009 Codebook. It shows each variable as its own row and provides the variable name, variable label, and coding values, including the reserved coding for missing data.

Slide 14 of 18

As a supplement to information contained in the PIRLS and TIMSS data files, system level information is provided by each participating education system and published in the PIRLS and TIMSS Encyclopedias. Each chapter focuses on one of the participating education systems, and provides a description of the policies and practices that guide that education system’s school organization, teacher education requirements, and the examinations and assessments used to measure achievement. The chapters in the PIRLS Encyclopedia emphasize language, literacy, and classroom reading instruction in the primary grades, while the TIMSS Encyclopedia chapters focus on the mathematics and science curricula and instruction in primary and secondary grades.

You can download the PIRLS and TIMSS Encyclopedias by clicking on the respective underlined screen text.

Slide 15 of 18

These modules have been designed to help you understand and analyze PIRLS, TIMSS, and PISA data. While completing these modules will provide a solid foundation of knowledge, the modules alone may not provide sufficient detail to ensure successful analyses that reflect a comprehensive understanding of the data and what they mean.

We suggest that you consult the user's manuals and other technical reports, which provide a wealth of information about how the different measures used in the study were developed, what they indicate, and how data from the measures are reported or provided in the data file.

Also, users may want to read through substantive or topical reports that use PIRLS, TIMSS, or PISA data to see what kind of information is available and what kinds of analyses can be conducted with the data. Those published by NCES can be found under "Publications and Products" at each study's homepage, and the NCES website includes a Bibliography Search Tool to search for journal articles using data from PIRLS, TIMSS, or PISA. Furthermore, creating tables and charts with the data through the International Data Explorer, or IDE, familiarizes users with the variables available. You can access the Bibliography Search Tool and IDE by clicking on the underlined screen text.

It is also important for users to review the data collection instruments that are available online so that they know exactly how questions were asked and understand the groups of people to or about whom questions were asked. As you review the instruments you should pay special attention to the paths through the questionnaire, including possible skip patterns.

Lastly, it is always a good idea to check the PIRLS, TIMSS, or PISA website occasionally for information about new data releases or issues that have been found in the data of which people should be aware.

Slide 16 of 18

The IEA IDB Analyzer is a stand-alone software program for analyzing data from all IEA surveys, including PIRLS and TIMSS, as well as OECD studies such as PISA. It can be used to merge together PIRLS or TIMSS data files that you need, as well as create SPSS code to analyze these data from complex sample designs with plausible values methodology, including the correct calculation of standard errors. The IDB Analyzer can be freely downloaded by clicking on the underlined screen text, and it will be discussed in-depth in the module titled 'Considerations for Analysis of IAP Data.'

Slide 17 of 18

You can obtain data from the IAP studies and related resources by clicking on the respective underlined screen text.

From the IEA–International Study Center website, click on either the “PIRLS” or “TIMSS” dropdown menus at the top of the screen and select a year. From there, click on the left tab “International Database” to download the data, which are available in SPSS and SAS formats, as well as useful resources such as the almanacs, codebooks, and user guides. Looking again at the left tabs, you can also access the assessment frameworks, international versions of the contextual questionnaires, international reports of the main findings, encyclopedias, and technical reports or online documentation of the methods and procedures.

To obtain PISA data from the OECD website, under “Results–Databases,” click on the year that you are interested in and then download the PISA database for that year. These data are available in SPSS and SAS formats. Here you can also download resources such as international versions of the contextual questionnaires, compendia, and codebooks. If you go back to the original screen, you can access international reports of the main findings under “Results–Key findings,” and assessment frameworks, data analysis manuals, and technical reports under “Methodology–manuals.”

Slide18 of 18

This module has provided you with an overview of the kinds of data files produced from the data collected for PIRLS, TIMSS, and PISA that are available for researchers to analyze. This module also described the resources that are available to learn more about the studies, the data, and the data files. The module’s objectives and the resources provided throughout the module are summarized here for your reference.

You may now proceed to the next module in the series, or click the exit button to return to the landing page.